THEORETICAL PERSPECTIVES ON LOCALISED KNOWLEDGE SPILLOVERS AND AGGLOMERATION*

Samuli Leppälä
Cardiff University†

August 2, 2016

Abstract

Empirical evidence shows that innovation is geographically concentrated, but whether localised knowledge spillovers provide a logically valid explanation for this phenomenon is unclear. I show that in the context of cost-reducing R&D spillovers between Cournot oligopolists the explanation is plausible: localised knowledge spillovers encourage agglomeration, but whether this leads to higher levels of effective R&D depends on the extent of the spillovers, the number of firms, and the industry's R&D efficiency. Contrary to the earlier theoretical work,

*I would like to thank David Collie, Huw Dixon, Marco Marinucci, Juuso Välimäki, two anonymous referees, and the seminar audiences at Cardiff University, The RSA Conference in Manchester, The Geography of Innovation Conference in Utrecht, and Industrial Organization: Theory, Empirics and Experiments Workshop in Alberobello for their helpful comments and suggestions. Any remaining errors are my own.

†Contact information: Cardiff University, Cardiff Business School, Colum Drive, Cardiff CF10 3EU, UK; Tel. +44(0)29 2087 6530; Email: leppalasm@cardiff.ac.uk

this illustrates a context, in which the localised knowledge spillover explanation is actualised, and provides necessary conditions, which can be empirically tested.

JEL classification: L13, O33, R32.

Keywords: agglomeration economies, innovation, R&D spillovers, location.

1 Introduction

Following the seminal research by Glaeser et al. (1992), there have been many empirical studies on the geography of innovation.[1] While some issues are still unclear, such as the role of industrial structure (Beaudry and Schiffauerova, 2009), there is definitive evidence that innovation is geographically concentrated (Asheim and Gertler, 2005; Feldman and Kogler, 2010). This outcome is typically attributed to localised knowledge spillovers, which are strongly bounded in space (Audretsch and Feldman, 2004; Döring and Schnellenbach, 2006). These spillovers are considered to induce agglomeration and thereby innovation in these locations. However, the empirical studies have been criticised due to the lack of a firm theoretical background and any direct evidence of spillovers (Breschi and Lissoni, 2001), which is problematic because there are various possible explanations for agglomeration economies (Rosenthal and Strange, 2004).

[1] One survey (Beaudry and Schiffauerova, 2009) and another meta-analysis (De Groot et al., 2015) considered 67 and 73 studies, respectively.

In contrast to empirical research, there has been little theoretical research into localised knowledge spillovers. Moreover, the theoretical basis of the localised knowledge spillover explanation of geographically concentrated innovation is far from definite. Since spillovers in the form of involuntary leakages to rivals increase the dissemination of R&D but also decrease the incentives to invest in it, the overall effect on both location choice and innovation is ambiguous. A few previous studies have considered endogenous, location-dependent spillovers in the context of Bertrand or Cournot competition (e.g. Van Long and Soubeyran, 1998; Baranes and Tropeano, 2003; Piga and Poyago-Theotoky, 2005), but they provide no clear support for the localised knowledge spillover explanation. Therefore, it needs to be addressed whether firms would choose to locate in close proximity in order to maximise spillovers, and if such agglomeration would lead to higher R&D levels and growth in output without explicit cooperation in R&D. As noted by several authors, among the micro-foundations of urban agglomeration, learning and knowledge spillovers are the least understood and there is urgent need for theoretical research, which informs empirical research rather than lags behind it (Duranton and Puga, 2004; Fujita and Krugman, 2004; Puga, 2010).

Localised knowledge spillovers can explain geographically concentrated innovation only if 1) localised spillovers promote agglomeration, and 2) agglomeration leads to a higher effective R&D output. If the first condition does not hold, then the explanation for agglomeration needs to be sought from other factors, such as labour market sorting or spinoff formation (Combes et al., 2008; Golman and Klepper, 2016). If the second condition does not hold, then the increase in innovation must come from elsewhere, such as labour mobility or R&D cooperation (Simonen and McCann, 2008). In this paper I present the first model that supports the localised knowledge spillover explanation. Furthermore, by identifying the conditions under which the explanation is sound it provides a way to empirically discriminate between the alternative explanations for geographically concentrated innovation.

In the model, I focus on output spillovers of cost-saving technology between non-cooperative Cournot oligopolists. Following empirical research, I assume that the extent of spillovers depends on the spatial proximity between firms. I do not consider the specific spillover mechanism, university-industry spillovers, or other agglomeration economies or diseconomies. The modelling choices are made in order to isolate the effect of localised knowledge spillovers and to demonstrate that circumstances where it leads to both agglomeration and more effective R&D exist. The results extend to other cases where firms choose R&D spillover rates non-cooperatively.

The model shows that agglomeration is always an equilibrium for any $n \geq 3$ firms, irrespective of the agglomeration spillover rate. However, agglomeration does not lead to higher effective R&D if the spillover rate is too high. The number of firms and R&D efficiency affect this threshold, which suggests that different industries benefit more than others from agglomeration. In addition to the agglomeration spillover rate, concentration, and R&D efficiency, empirical research should pay more attention to the type of spillovers and the employed R&D proxies.

The rest of this paper is organized as follows. Section 2 reviews the basic models of knowledge spillovers and their extensions to the context of location choice. Based on these insights, Section 3 presents a Cournot oligopoly model, which provides a suitable framework for examining the localised knowledge spillover explanation. Finally, I summarise the results in Section 4 and provide suggestions for future research.

2 Literature Review

Since the research on localised knowledge spillovers and agglomeration has been largely empirically orientated, the theoretical basis must be sought from elsewhere.[2] As Iammarino and McCann (2006, 1024) note, "understanding the reasons why particular observed clusters exist requires a careful consideration of central issues in industrial organisation". There have been numerous studies of knowledge spillovers in the industrial organisation literature (De Bondt, 1997; Sena, 2004). Two seminal studies on the effect of knowledge spillovers on R&D incentives are d'Aspremont and Jacquemin (1988) and Kamien et al. (1992), in which the R&D spillovers are formalised by the

[2] Reviews of the empirical literature are provided in Breschi and Lissoni (2001); Asheim and Gertler (2005); Döring and Schnellenbach (2006); Beaudry and Schiffauerova (2009); Feldman and Kogler (2010); De Groot et al. (2015).

effective R&D equation:

$$X_i = x_i + \beta \sum_{j \neq i} x_j,$$

where $\beta \in [0,1]$ is the spillover rate between firms.[3] x_i in d'Aspremont and Jacquemin (1988) is firm i's own R&D output, which together with the spillovers from other firms forms its effective R&D output, X_i. In Kamien et al. (1992), however, these are investments in R&D, i.e., self-financed and effective, respectively. The R&D output in both cases is typically considered to be a cost reduction, but the logic is similar for quality-enhancing R&D.

Despite the similarities, the outcomes of these models differ in some relevant respects and previous studies have discussed their relative merits (see Amir, 2000; Amir et al., 2008). For example, Amir (2000) considered the additive spillovers of the output spillover model to be less realistic, but noted that these might be appropriate in some cases, especially when modelling agglomeration economies that assume additive benefits. One way of understanding the difference between these two processes is whether the firms jointly refine the same technology or if they develop different but additive technologies.[4]

The implications of the models for empirical research are as follows. Both models predict that the firms' R&D efforts decrease in the spillover rate β.

[3]For a discussion of other functional forms and their implications, see Knott et al. (2009).

[4]Note that, by construction, both models ignore the risk of duplication.

The same holds for the the effective R&D, X_i, if the spillovers are in inputs, whereas the maximum effective R&D is reached when $\beta = 0.5$ in the case of output spillovers.[5] As such, some degree of imperfection in the spillovers is important. Therefore, it should be important whether the R&D variable employed measures a firm's self-financed or effective R&D. Furthermore, input spillovers are not likely to explain the spatial concentration of innovation because after controlling for other influences higher spillovers always lead to less R&D, however measured.

Whether the spillovers occur in R&D outputs or inputs may vary between industries and this is ultimately an empirical question concerning the additivity of different inventions. However, my aim is only to test whether the proposed explanation is logically valid. Therefore, my model employs the approach of d'Aspremont and Jacquemin (1988) because it is the most favourable of the two cases. However, we can expect the agglomeration spillover rate to play a critical role. Similarly, I utilise a Cournot model, because spillovers can only decrease the R&D levels in a homogeneous good price competition.

The spillover rate creates opposing efficiency and incentive effects and previous studies have shown that while firms prefer to minimise leakage to their rivals unilaterally, they would choose extremal spillovers cooperatively (Poyago-Theotoky, 1999; Amir et al., 2003). This makes it interesting to see how spillovers affect non-cooperative location choices. One way of combin-

[5] De Bondt et al. (1992) proved that the latter result holds for any number of firms n.

ing knowledge spillovers with location choice is to introduce them into the Hotelling model. A few studies have employed this approach but the equilibrium outcomes varied between agglomeration and dispersion, depending on the model characteristics (Mai and Peng, 1999; Piga and Poyago-Theotoky, 2005; Li and Zhang, 2013). However, the Hotelling model has two unwanted characteristics given our aims. First, a linear city is not likely to be a suitable model for representing the empirical findings on the differences between peripheries and cities. Second, in the Hotelling model, the location choice also affects the market shares as well as the spillover rate. There are certainly centripetal and centrifugal forces other than localised knowledge spillovers (see Fujita and Krugman, 2004; Fujita and Thisse, 2013), but the aim here is to establish whether it is logically true that these spillovers provide a reason to agglomerate in the absence of interfering factors.

With respect to non-Hotelling approaches, some previous studies considered whether two firms choosing to locate in the same region might lead to spillovers. In Alsleben (2005), labour poaching, which occurs when firms are located in the same region, leads firms to choose dispersion. By contrast, in Baranes and Tropeano (2003), selecting the same location created more competition and made the firms more willing to share their R&D. However, knowledge that is fully embodied in human capital or the voluntary sharing of R&D does not conform with the strict meaning of knowledge spillovers because no externalities are involved (Breschi and Lissoni, 2001). Other existing papers that employ a similar discrete location choice approach are likewise limited in addressing the localised knowledge spillover explanation of spatially concentrated innovation by concentrating on the location choice and keeping the R&D investment decisions exogenous (Gersbach and Schmutzler, 1999; Fosfuri and Rønde, 2004; Combes and Duranton, 2006).

In three papers, which are most closely related to my approach, the firms choose the distance (or the level of technological differentiation) between each other, which then determines the spillover rates (but not other market factors). In Van Long and Soubeyran (1998), three firms choose to agglomerate given any level of R&D investments when the (input) spillover effect is convex in distance. However, the study does not reveal how agglomeration affects the R&D levels, which could also affect the location choices subsequently. In Gil Moltó et al. (2005), a duopoly model is employed where the R&D levels are endogenous.[6] The study shows that firms maximise or almost maximise the spillovers depending on the R&D efficiency and the highest attainable spillover rate. In addition, it was shown that the spillover choice leads to a decrease in R&D propensity, x_i, but the impact on the effective R&D, X_i, was not considered.

Importantly, a duopoly model is unable to determine whether there is an agglomeration equilibrium with more than two firms where one firm's decision to deviate does not affect the spillover rates between the other firms. A three-firm oligopoly is considered in Mota and Brandão (2004), but un-

[6]While the endogenous (output) spillover rate depends on technological differentiation in their model, the situation is similar if this is caused by the distance between the firms.

fortunately their simplifying assumption of the identical equilibrium R&D levels is not appropriate if the (output) spillover rates are not identical as well. By considering this issue and extending the baseline model to an n-firm Cournot oligopoly, we can study whether agglomeration is an equilibrium and the circumstances under which it also leads to higher effective R&D. The previous theoretical models have either been inconsistent with this outcome or have not studied both aspects of the issue. In contrast, the following model provides an example where the localised knowledge spillover explanation is plausible together with necessary conditions for empirical testing.

3 The Model

In this section, we consider a three-stage game between $n \geq 3$ Cournot oligopolists, which produce a homogeneous output. The inverse demand function is given by $P = a - Q$, where $Q = \sum_{i=1}^{n} q_i$ is the total quantity produced and $a > Q \geq 0$. The unit cost of firm $i, i = 1, ..., n$, is $c_i = c - X_i$, where c is the initial marginal cost, X_i is the effective cost reduction due to R&D, and $a > c > X_i \geq 0$. Hence, I assume that marginal costs are always positive and that some production takes always place.

A firm's effective R&D output is

$$X_i = x_i + \sum_{j \neq i} \beta(d_{ij}) x_j.$$

As in d'Aspremont and Jacquemin (1988), the cost of the firm's own R&D output x_i is quadratic and it is given by

$$C(x_i) = \frac{1}{2}\gamma x_i^2,$$

where $\gamma > 0$ is an inverse measure of the efficiency of R&D. In particular, γ is the slope of the marginal R&D cost curve and reflects the R&D cost efficiency of the industry: the higher the γ, the more expensive it is to reduce costs or increase quality.

We do not explicitly consider the exact spillover mechanism, but instead follow the empirical research in assuming that the spillovers are simply decreasing in distance. The interest here is on whether, in the absence of any interfering factors, the consequences of the assumption that the firms can affect the spillover rate through their choice of location are as presumed. Hence, the output spillovers from other firms depend on the spillover rate $\beta(d_{ij})$, which is a positive and decreasing function of the geographic distance d_{ij} between firms i and j $(i \neq j)$, i.e.,

$$0 \leq \beta(d_{ij}) \leq \bar{\beta} \leq 1,$$

and $\beta'(d_{ij}) < 0$ and $\beta(0) = \bar{\beta}$. For convenience, we denote this by $\beta_{ij} = \beta(d_{ij})$ and we concentrate on the choice of β while keeping in mind the assumption that it is chosen indirectly through the choice of distance.

The agglomeration spillover rate $\bar{\beta}$ is the upper bound that can be achieved by setting the distance to zero and choosing the same location.[7] This maximal spillover rate can be limited by other factors, such as labour mobility, technological (dis)similarity, or intellectual property rights.[8] Similarly, there could also be a lower bound to localised knowledge spillovers, but this is not our concern because we concentrate on the agglomeration case. I assume that the transportation costs and any other costs related directly to the location choice are zero, thereby allowing us to focus on how localised knowledge spillovers alone affect the location choice. This implies that the results can also be extended to other cases of endogenous knowledge spillovers (c.f. Gil Moltó et al., 2005).

The timing of the three-stage game is as follows.

1. The firms choose their distance d_{ij} from each other and hence the spillover rate β_{ij} between them.

2. The firms choose their own cost reduction levels, x_i.

3. The firms choose their output levels, q_i, via Cournot competition.

In each stage, the choices are made simultaneously and discounting between the stages is ignored for simplicity. We solve the game by backward

[7] Even co-located firms may, of course, have asymmetric spillover rates, which would typically result from other differences in their relative positions (De Bondt and Henriques, 1995; Vandekerckhove and De Bondt, 2008). However, differences in the realised spillovers need not imply that the firms expect their spillover rates to be asymmetric.

[8] The spillover rate may further depend on the firm's own R&D effort. This absorptive capacity effect has been shown to increase the firms' R&D efforts, but leaving the qualitative results of the basic model unchanged (Martin, 2002). As such, it is not expected that absorptive capacity would change the way how localised knowledge spillovers affect location choice and innovation in this model.

induction to determine whether agglomeration can be a Nash equilibrium, which also maximises the firm's effective R&D. The previous literature alludes that both outcomes are dependent on $\bar{\beta}$. Since we do not consider whether other equilibria exist, I do not need to make any explicit assumptions regarding the location space, except that there is at least one dimension, or the concavity of spillovers in space. Without loss of generality, I will assume that all the other firms except i are agglomerated and we concentrate on firm i's location choice.[9] Thus, if $d_{jk} = 0, \forall j, k \in \{n - i\}, j \neq k$, which implies $\beta_{jk} = \bar{\beta}$ and $\beta_{ij} = \beta_{ik} = \beta$, then we determine the necessary conditions for $\beta = \bar{\beta}$ to maximise firm i's effective R&D and profit.

3.1 Production Stage

In the production stage, firm i maximises its profit function, which is given by

$$\pi_i = (a - Q - c_i)q_i.$$

The Cournot equilibrium output is

$$q_i^* = \frac{a - nc_i + \sum_{j \neq i} c_j}{n+1} = \frac{a - c + nX_i - \sum_{j \neq i} X_j}{n+1} \tag{1}$$

[9]This relates to the analysis of the "sustain points" in the new economic geography literature (Fujita et al., 1999).

for all firms $i \in n$. The total industry output is

$$Q = \frac{n(a-c) + \sum_{i=1}^{n} X_i}{n+1}$$

and the consumer surplus is $CS = \frac{1}{2}Q^2$. As expected, there is a positive effect of R&D on the economic activity and welfare since $\partial Q/\partial X_i > 0$.

3.2 R&D Investment Stage

In stage 2, the firms choose their R&D levels. Given the subsequent output levels, firm i chooses x_i in order to maximise

$$\pi_i = (q_i^*)^2 - \frac{1}{2}\gamma x_i^2,$$

where q_i^* is given by equation (1). Assuming that firms other than i are agglomerated, $d_{jk} = 0$, implies $\beta_{jk} = \bar{\beta}$ and $\beta_{ij} = \beta_{ik} = \beta$. The first order condition gives the best response function

$$x_i(x_j) = \frac{2(a - c + (n\beta - (n-2)\bar{\beta} - 1)\sum_{j \neq i} x_j)(n - (n-1)\beta)}{\gamma(n+1)^2 - 2(n - (n-1)\beta)^2} \tag{2}$$

for firm i. This shows us that the R&D outputs x_j are strategic substitutes for x_i if $n\beta - (n-2)\bar{\beta} - 1 < 0$ and complements if the inequality is reversed.

Similarly, the best response function for the other firms is

$$x_j(x_i, x_k) = \frac{2(a - c + (2\beta - 1)x_i + (3\bar{\beta} - \beta - 1)\sum x_k)(n - \beta - (n-2)\bar{\beta})}{\gamma(n+1)^2 - 2(n - \beta - (n-2)\bar{\beta})^2} \tag{3}$$

$\forall j, k \in \{n - i\}, j \neq k$. Hence, the R&D output x_i is a strategic substitute for x_j if $\beta < 1/2$, and x_j and x_k are strategic substitutes for each other if $3\bar{\beta} - \beta - 1 < 0$.

The second order conditions in the R&D stage require that the numerators in the best response functions are positive. This holds for all $\beta, \bar{\beta} \in [0, 1]$ when $\gamma > 2n^2/(n+1)^2$. The stability condition requires that the best response functions cross correctly (Henriques, 1990), which holds for all $\beta, \bar{\beta} \in [0, 1]$ when $\gamma > 2n(2n-1)/(n+1)^2$.

I assume that firms $j \neq i$ make a symmetric choice: x_{-i}. Then, using the best response functions (2) and (3), we obtain the following equilibrium R&D output levels:

$$x_i^* = 2(a-c)(n - (n-1)\beta)\frac{A}{C} \tag{4}$$

and

$$x_{-i}^* = 2(a-c)(n - \beta - (n-2)\bar{\beta})\frac{D}{C}, \tag{5}$$

where

$$A = (n+1)\gamma - 2(\bar{\beta} - 1)(\beta - \bar{\beta})n^2 + ((8\beta + 6)\bar{\beta} - 8\bar{\beta}^2 - 2\beta^2 - 2\beta - 2)n$$

$$+2\beta + 8\bar{\beta}^2 - (8\beta + 4)\bar{\beta} + 2\beta^2,$$

$$C = (8n - 4n^2 - 4)\beta^4 + ((16\bar{\beta} - 4)n^2 - 20\bar{\beta}n + (4 - 4\bar{\beta})n^3 + 8\bar{\beta})\beta^3$$

$$+((2n - 2n^3)\gamma + 8\bar{\beta} - 4 + (4 - 4\bar{\beta})n + (4 - 8\bar{\beta})n^2 + (4\bar{\beta} - 4)n^3)\beta^2$$

$$+((12\bar{\beta} - 20\bar{\beta}^2 - 4)n^2 + ((6 - 2\bar{\beta})n^3 - 12\bar{\beta} + 4 + 8\bar{\beta}n^2 + (2 - 2\bar{\beta})n)\gamma$$

$$+(4\bar{\beta}^2 - 4\bar{\beta})n^3 - 16\bar{\beta}^2 + 8\bar{\beta} + (32\bar{\beta}^2 - 12\bar{\beta})n)\beta + (n^3 + 3n^2 + 3n + 1)\gamma^2$$

$$+((4\bar{\beta}^2 - 4\bar{\beta} - 2)n^3 + 16\bar{\beta}^2 + (8\bar{\beta} - 12\bar{\beta}^2 - 6)n^2 + (4\bar{\beta} - 4)n - 8\bar{\beta})\gamma$$

$$+(16\bar{\beta}^2 - 12\bar{\beta} + 4)n^2 + (4\bar{\beta} - 4\bar{\beta}^2)n^3 + (8\bar{\beta} - 16\bar{\beta}^2)n,$$

and

$$D = (n+1)\gamma - 2n + (2 - 2n)\beta^2 + (4n - 2)\beta.$$

The interior and positive solutions for the R&D outputs, particularly that $A > 0$, are guaranteed for $\gamma > (n+1)/2, \forall \beta, \bar{\beta} \in [0,1]$.[10] I make the following assumption:

Assumption 1 $\gamma > (n+1)/2$.

With Assumption 1, I limit the analysis to positive, locally stable equilibrium values in the R&D investment stage, ruling out either maximal R&D investments or corner solutions.[11]

The equilibrium R&D outputs (4) and (5) yield effective cost reductions

[10] Further details available on request.

[11] Note that corner solutions could be relevant for studying whether any firm would ever completely isolate itself.

$X_i = x_i^* + \beta(n-1)x_{-i}^*$ and $X_{-i} = (1+\bar{\beta}(n-2))x_{-i}^* + \beta x_i^*$. The reader should keep in mind that here β is a function. Now, we proceed to the first main result.

Proposition 1 *Agglomeration leads to higher effective R&D only if the agglomeration spillovers are moderate, i.e.,* $\bar{\beta} \leq \hat{\beta}$*, where* $\hat{\beta} \in \left(\frac{n-1}{n+1}, \frac{n-1}{n}\right)$*,* $\partial\hat{\beta}/\partial\gamma > 0$*.*

Proof. See Appendix A.

Proposition 1 shows that for low to medium spillovers, firm i's effective R&D can be higher if it is agglomerated with the other firms. However, there is an inverted-U relationship between innovation and agglomeration spillovers. If the agglomeration spillovers are high, the firm could increase its effective R&D by not agglomerating with the other firms. In this case, the increase in own R&D output would compensate for the lower spillovers. Since this effect depends on the number of firms, n, this provides a moving window for the critical spillover rate. With a larger number of firms, the higher rate of agglomeration spillovers would still increase the firm's effective R&D. Intuitively, n affects both the quantity of spillovers that a firm can enjoy as well as the degree of strategic effect that its agglomeration decision has on the R&D choices of the other firms. As such, the critical spillover rate is higher for $n \geq 3$ firms than in the standard model (c.f. De Bondt et al., 1992). However, the meaning of $\hat{\beta}$ is different. Instead of measuring the common spillover rate that maximises each firm's effective R&D, it provides an important counterfactual condition, i.e., the spillover rate beyond which a firm would enjoy higher effective R&D outside the agglomeration. Thus, the prediction still remains that the highest effective R&D when comparing across different agglomerations is gained when the agglomeration spillover rates are exactly intermediate.

To some extent, the effect of agglomeration spillovers depends on R&D cost efficiency, which determines the critical spillover rate within the bounds. A larger γ moves the critical rate closer to the upper bound, in which case, higher agglomeration spillovers increase effective R&D due to cost savings. However, the magnitude of this effect is small, which is partly a consequence of Assumption 1. By approximating the bounds for some values of n:

$$n = 3 \rightarrow \hat{\beta} \in (0.6498, 2/3)\,, n = 5 \rightarrow \hat{\beta} \in (0.7796, 0.8)$$

$$n = 10 \rightarrow \hat{\beta} \in (0.8930, 0.9)\,, n = 25 \rightarrow \hat{\beta} \in (0.9592, 0.96)\,,$$

we can see that they tend to be very close to each other. γ only has a small effect, so we can say that the critical spillover rate occurs slightly before $(n-1)/n$ in general.

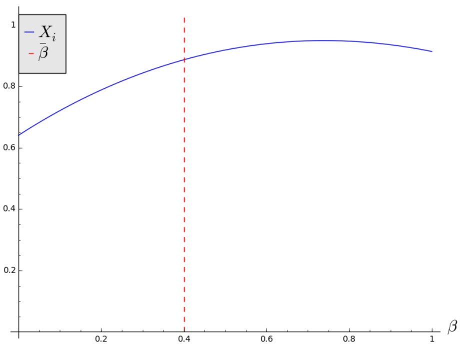
Note that not exceeding the critical spillover rate is only a necessary condition for agglomeration to maximise effective R&D. The sufficient complexity of the effective R&D function makes it infeasible to prove that in this case agglomeration also ensues the global maximum. However, based on my numerical computations, it strongly seems to be the case that X_i is concave in β over the relevant interval as illustrated in Figure 1.[12] Be that as it may, the main purpose of this paper is to identify necessary conditions of the localised knowledge spillover explanation, which can then be empirically tested. One can further observe from Figure 1 how, as the agglomeration spillover rate, $\bar{\beta}$, increases and moves right, the peak of X_i moves left. The peak is left to $\bar{\beta}$ when the latter gets sufficiently high, but a higher n requires an even higher $\bar{\beta}$ for this to take place.

3.3 Location Choice Stage

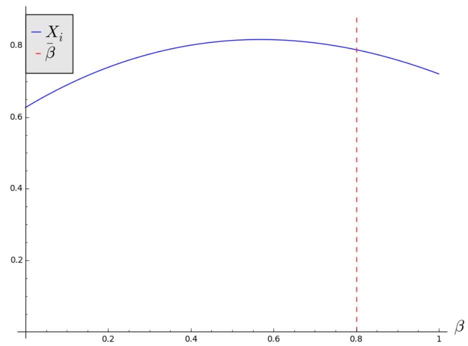
Although moderate agglomeration spillovers imply higher effective R&D, the final step is to check the range of agglomeration spillovers $\bar{\beta}$ for which agglomeration can be an equilibrium outcome. If the range of spillovers does not overlap with Proposition 1, then the localised knowledge spillover explanation is not a logically valid even within this model. As before, I assume that all the other firms are agglomerated and we concentrate on firm i's decision. Given the anticipated outcome of stages 2 and 3, and the equilibrium cost reductions (4) and (5), firm i's profit function in stage 1 is now

$$\begin{aligned}\pi_i &= \frac{(a-c+nX_i-(n-1)X_{-i})^2}{(n+1)^2} - \frac{1}{2}\gamma(x_i^*)^2 \\ &= \frac{(a-c+(n-(n-1)\beta)x_i^*+(n-1)(n\beta-(n-2)\bar{\beta}-1)x_{-i}^*)^2}{(n+1)^2} - \frac{1}{2}\gamma(x_i^*)^2.\end{aligned} \tag{6}$$

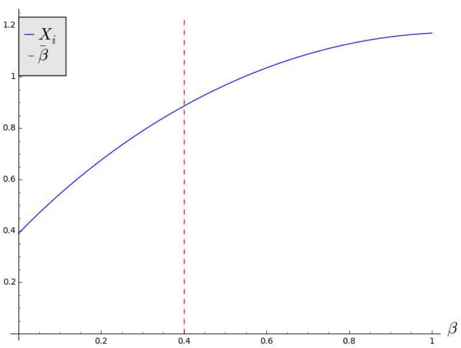
[12] These patterns are robust to a wide range of variations in parameter values.



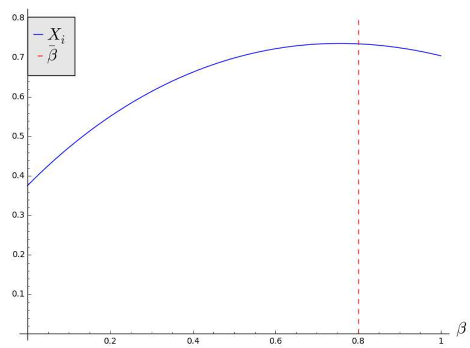
(a) $n = 3$ and $\bar{\beta} = 0.4$



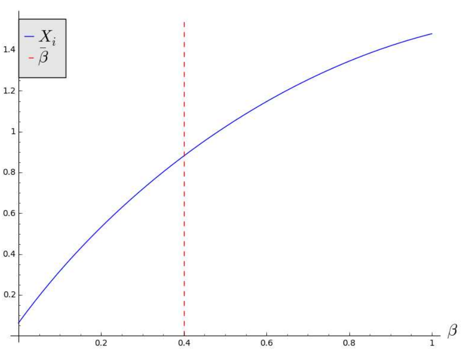
(b) $n = 3$ and $\bar{\beta} = 0.8$



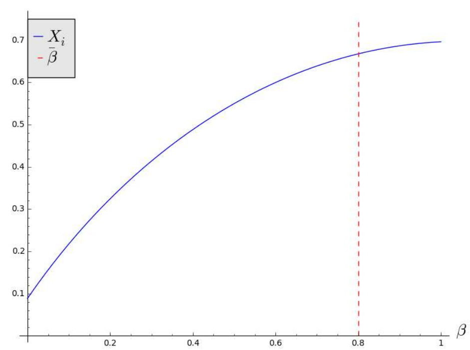
(c) $n = 5$ and $\bar{\beta} = 0.4$



(d) $n = 5$ and $\bar{\beta} = 0.8$



(e) $n = 10$ and $\bar{\beta} = 0.4$



(f) $n = 10$ and $\bar{\beta} = 0.8$

Figure 1: Effective R&D and the spillover rate (with $a - c = 10$, $\gamma = 6$, and different ns and $\bar{\beta}$s).

This brings us to the final proposition.

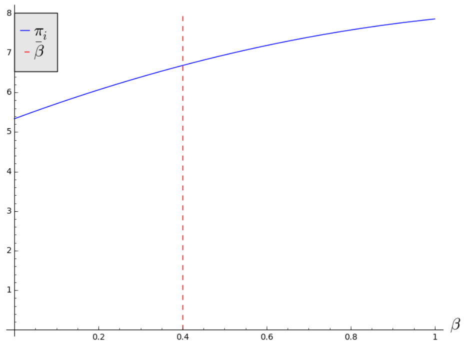
Proposition 2 *Agglomeration is a possible equilibrium outcome for n firms given any rate of agglomeration spillovers $\bar{\beta}$.*

Proof. See Appendix A.

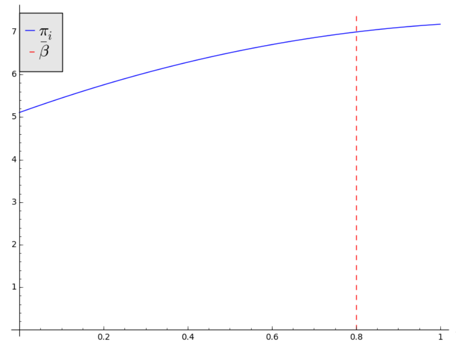
Proposition 2 means that agglomeration is always a possible outcome, irrespective of the spillover rate it yields. In the absence of any offsetting factors[13], this holds for any $n \geq 3$ firms (cf. Gil Moltó et al., 2005). As the agglomeration outcome is not ruled out by any $\bar{\beta}$ in the model, the necessary condition as identified in Proposition 1 can be used for testing whether localised knowledge spillovers explain the spatial concentration of innovation. Note that Proposition 2 claims only that agglomeration maximises the profit at least locally. Again, the functional complexity of X_i and π_i makes it infeasible to prove that this is also the global maximum. My numerical computations strongly suggest that this is the case, however. Looking at Figure 2, it seems that an increase in n, ceteris paribus, changes the $\pi_i(\beta)$ function from concave to convex. In both cases, however, π_i is strictly increasing in β (hence, decreasing in d_{ij}) over the relevant interval.

Proposition 2 implies that despite the mixed effects on the effective R&D, localised knowledge spillovers create a centripetal force for 3 or more firms.

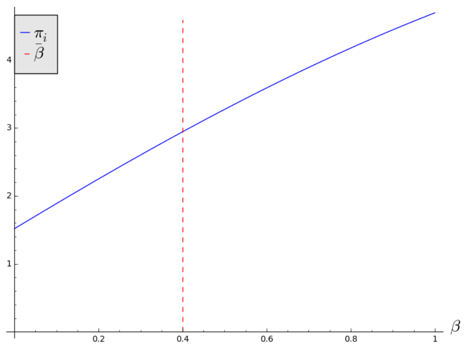
[13] In addition to any agglomeration diseconomies, asymmetric spillover rates might affect the outcome. If the spillover rates are different enough, we could have a separating equilibrium where only firms with high incoming and low outgoing spillover rates agglomerate (see Livanis and Lamin, 2016). Whether this outcome maximises the effective R&D of both agglomerated and isolated firms is, nevertheless, likely to depend on how their weighted average spillover rate compares to the critical rate of Proposition 1.



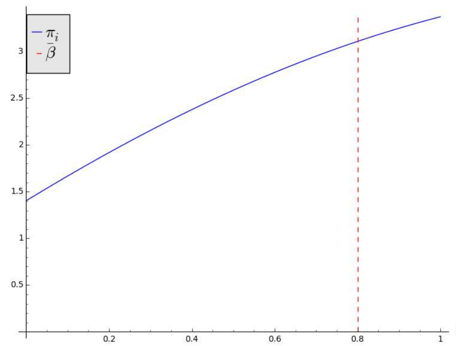
(a) $n = 3$ and $\bar{\beta} = 0.4$



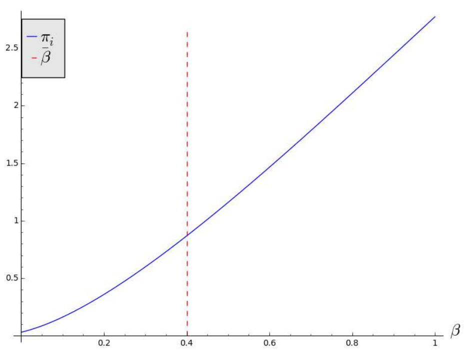
(b) $n = 3$ and $\bar{\beta} = 0.8$



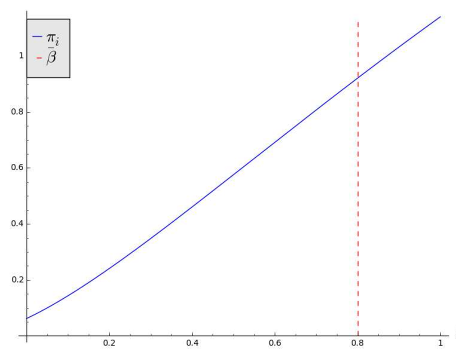
(c) $n = 5$ and $\bar{\beta} = 0.4$



(d) $n = 5$ and $\bar{\beta} = 0.8$



(e) $n = 10$ and $\bar{\beta} = 0.4$



(f) $n = 10$ and $\bar{\beta} = 0.8$

Figure 2: Profit and the spillover rate (with $a - c = 10$, $\gamma = 6$, and different ns and $\bar{\beta}$s).

While a firm wants to minimise the R&D leaked to its rivals, the incentive to similarly free-ride on their efforts is stronger. It is intuitive that a firm prefers to agglomerate when this implies higher effective R&D, but it is less obvious when this does not occur. However, locating outside the agglomeration would imply less spillovers and more own R&D investment, and hence a lower profit.

Naturally, the presence of congestion or transportation costs, or any other centripetal or centrifugal forces, would affect the location decision in reality as well. As such, this suggests that when firms are observed to be dispersed this is due to stronger centrifugal forces despite of localised knowledge spillovers. However, Proposition 1 then provides the counterfactual for assessing whether the dispersed firms would have been more innovative if agglomerated or the agglomerated firms when dispersed. Therefore, the agglomeration spillover rate, number of firms, and R&D cost efficiency play important roles in determining whether localised knowledge spillovers explain the spatial concentration of innovation. Furthermore, these factors may help to explain the observed differences between industries (Döring and Schnellenbach, 2006; De Groot et al., 2015).

4 Conclusion

The standard explanation for the pattern of geographically concentrated innovation has been localised knowledge spillovers. In the present study, I analysed the theoretical validity of this explanation in the context of Cournot oligopolists. Indeed, it holds that agglomeration is a possible equilibrium outcome of this model. However, it is not always the case that agglomeration will also imply higher effective R&D for these firms. Hence, localised knowledge spillovers may explain geographical concentration of innovation, but certain conditions still need to be met.

The implications from the theory to empirical research are as follows. Based on the review of the earlier literature, we conjecture that whether knowledge spillovers take place in R&D inputs or outputs is likely to be critical. That is, it is expected that only in industries, in which the technological space is characterised by additive inventions, can the relationship between localised knowledge spillovers and innovation be positive. Furthermore, it can be of great importance whether the employed R&D proxy measures expenditures, own R&D output, or the effective R&D, which also includes spillovers. Whether these respond similarly or differently to changes in the spilllover rate may further indicate if the industry is characterised by output or input spillovers.

The output spillover model analysed in this paper predicts that localised knowledge spillovers do facilitate agglomeration but that the relationship with innovation is non-linear. After controlling for other factors, such as the market size, we would expect agglomerated firms to be more innovative as long as the agglomeration spillovers are not too high. Furthermore, the critical spillover rate is conditional on the number of firms and R&D efficiency such that a higher spillover rate is advantageous if there are more firms or the industry's R&D activities are more costly to perform. All of these factors are likely to vary among different industries and technologies, and thus they provide interesting hypotheses for testing in subsequent empirical research.

By isolating localised knowledge spillovers from other interfering factors, I have provided the first theoretical model that may be able to explain the spatial concentration of innovation. Naturally, there might be other model specifications with different functional forms, competitive settings, or spillover mechanisms such as networks, where the explanation can be found to hold as well. Further theoretical research could also consider the existence of other equilibria, the impact of R&D cooperation complemented with full welfare analysis, or inter-industry spillovers, which were not addressed here. Hence, demand for careful theoretical work as well as theoretically grounded empirical studies exists.

References

Alsleben, C. (2005). The downside of knowledge spillovers: An explanation for the dispersion of high-tech industries. *Journal of Economics*, 84(3):217–248.

Amir, R. (2000). Modelling imperfectly appropriable R&D via spillovers. *International Journal of Industrial Organization*, 18(7):1013–1032.

Amir, R., Evstigneev, I., and Wooders, J. (2003). Noncooperative versus cooperative R&D with endogenous spillover rates. *Games and Economic Behavior*, 42(2):183–207.

Amir, R., Jin, J. Y., and Troege, M. (2008). On additive spillovers and returns to scale in R&D. *International Journal of Industrial Organization*, 26(3):695–703.

Asheim, B. and Gertler, M. (2005). The geography of innovation. In Fagerberg, J., Mowery, D., and Nelson, R., editors, *The Oxford Handbook of Innovation*, pages 291–317. Oxford University Press Oxford, Oxford.

Audretsch, D. B. and Feldman, M. P. (2004). Knowledge spillovers and the geography of innovation. In Henderson, J. V. and Thisse, J.-F., editors, *Handbook of Regional and Urban Economics*, volume 4, pages 2713–2739. Elsevier.

Baranes, E. and Tropeano, J.-P. (2003). Why are technological spillovers spatially bounded? A market orientated approach. *Regional Science and Urban Economics*, 33(4):445–466.

Beaudry, C. and Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, 38(2):318–337.

Breschi, S. and Lissoni, F. (2001). Knowledge spillovers and local innovation systems: a critical survey. *Industrial and Corporate Change*, 10(4):975–1005.

Combes, P.-P. and Duranton, G. (2006). Labour pooling, labour poaching, and spatial clustering. *Regional Science and Urban Economics*, 36(1):1–28.

Combes, P.-P., Duranton, G., and Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2):723–742.

d'Aspremont, C. and Jacquemin, A. (1988). Cooperative and noncooperative R&D in duopoly with spillovers. *American Economic Review*, 78(5):1133–1137.

De Bondt, R. (1997). Spillovers and innovative activities. *International Journal of Industrial Organization*, 15(1):1–28.

De Bondt, R. and Henriques, I. (1995). Strategic investment with asymmetric spillovers. *Canadian Journal of Economics*, 28(3):656–74.

De Bondt, R., Slaets, P., and Cassiman, B. (1992). The degree of spillovers and the number of rivals for maximum effective R&D. *International Journal of Industrial Organization*, 10(1):35–54.

De Groot, H. L., Poot, J., and Smit, M. J. (2015). Which agglomeration externalities matter most and why? *Journal of Economic Surveys*, forthcoming.

Döring, T. and Schnellenbach, J. (2006). What do we know about geographical knowledge spillovers and regional growth?: a survey of the literature. *Regional Studies*, 40(3):375–395.

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies. In Henderson, J. V. and Thisse, J.-F., editors, *Handbook of Regional and Urban Economics*, volume 4, pages 2063–2117. Elsevier.

Feldman, M. P. and Kogler, D. F. (2010). Stylized facts in the geography of innovation. In Hall, B. H. and Rosenberg, N., editors, *Handbook of The Economics of Innovation, Vol. 1*, pages 381–410. North-Holland.

Fosfuri, A. and Rønde, T. (2004). High-tech clusters, technology spillovers, and trade secret laws. *International Journal of Industrial Organization*, 22(1):45–65.

Fujita, M. and Krugman, P. (2004). The new economic geography: Past, present and the future. *Papers in Regional Science*, 83(1):139–164.

Fujita, M., Krugman, P. R., and Venables, A. (1999). *The Spatial Economy: Cities, Regions, and International Trade*. MIT press, Cambridge, MA.

Fujita, M. and Thisse, J.-F. (2013). *Economics of Agglomeration: Cities, Industrial Location, and Globalization*. Cambridge University Press.

Gersbach, H. and Schmutzler, A. (1999). External spillovers, internal spillovers and the geography of production and innovation. *Regional Science and Urban Economics*, 29(6):679–696.

Gil Moltó, M., Georgantzís, N., and Orts, V. (2005). Cooperative R&D with endogenous technology differentiation. *Journal of Economics & Management Strategy*, 14(2):461–476.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, 100(6):1126–1152.

Golman, R. and Klepper, S. (2016). Spinoffs and clustering. *RAND Journal of Economics*, 47(2):341–365.

Henriques, I. (1990). Cooperative and noncooperative R&D in duopoly with spillovers: Comment. *American Economic Review*, 80(3):638–40.

Iammarino, S. and McCann, P. (2006). The structure and evolution of industrial clusters: Transactions, technology and knowledge spillovers. *Research Policy*, 35(7):1018–1036.

Kamien, M. I., Muller, E., and Zang, I. (1992). Research joint ventures and R&D cartels. *American Economic Review*, 82(5):1293–1306.

Knott, A. M., Posen, H. E., and Wu, B. (2009). Spillover asymmetry and why it matters. *Management Science*, 55(3):373–388.

Li, C. and Zhang, J. (2013). R&D competition in a spatial model with technical risk. *Papers in Regional Science*, 92(3):667–682.

Livanis, G. and Lamin, A. (2016). Knowledge, proximity and R&D exodus. *Research Policy*, 45(1):8–26.

Mai, C. and Peng, S. (1999). Cooperation vs. competition in a spatial model. *Regional Science and Urban Economics*, 29(4):463–472.

Martin, S. (2002). Spillovers, appropriability, and R&D. *Journal of Economics*, 75(1):1–32.

Mota, I. and Brandão, A. (2004). Firms' location and R&D cooperation in an oligopoly with spillovers. *Revista Portuguesa de Estudos Regionais*, 6:183–200.

Piga, C. and Poyago-Theotoky, J. (2005). Endogenous R&D spillovers and locational choice. *Regional Science and Urban Economics*, 35(2):127–139.

Poyago-Theotoky, J. (1999). A note on endogenous spillovers in a non-tournament R&D duopoly. *Review of Industrial Organization*, 15(3):253–262.

Puga, D. (2010). The magnitude and causes of agglomeration economies. *Journal of Regional Science*, 50(1):203–219.

Rosenthal, S. S. and Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In Henderson, J. V. and Thisse, J.-F., editors, *Handbook of Regional and Urban Economics*, volume 4, pages 2119–2171. Elsevier.

Sena, V. (2004). The return of the prince of Denmark: A survey on recent developments in the economics of innovation. *The Economic Journal*, 114(496):F312–F332.

Simonen, J. and McCann, P. (2008). Firm innovation: The influence of r&d

cooperation and the geography of human capital inputs. *Journal of Urban Economics*, 64(1):146–154.

Ştefănesu, D., Gerdt, V., and Yevlakov, S. (2010). Estimations of positive roots of polynomials. *Journal of Mathematical Sciences*, 168(3):468–474.

Van Long, N. and Soubeyran, A. (1998). R&D spillovers and location choice under Cournot rivalry. *Pacific Economic Review*, 3(2):105–119.

Vandekerckhove, J. and De Bondt, R. (2008). Asymmetric spillovers and investments in research and development of leaders and followers. *Economics of Innovation and New Technology*, 17(5):417–433.

Appendix A

Several parts of the proofs of Propositions 1 and 2 rely on the positivity or negativity of polynomial functions of n. For convenience, we establish these collectively in the following lemma.

Lemma 1 *The signs of the polynomials are as follows.*

i $P_1(n) = -3n^5 + 19n^4 + 23n^3 + 133n^2 + 132n + 36 < 0$ *if* $n > 8$.

ii $P_2(n) = n^5 + 5n^4 - 26n^3 + 178n^2 - 287n + 81 > 0$ *if* $n \geq 3$.

iii $P_3(n) = n^3 - n^2 - 8n + 24 > 0$ *if* $n \geq 3$.

iv $P_4(n) = n^5 + 17n^4 - 51n^3 + 61n^2 + 10n - 6 > 0$ *if* $n \geq 3$.

v $P_5(n) = n^5 + n^4 - 9n^3 - 17n^2 + 128n - 72 > 0$ *if* $n \geq 3$.

vi $P_6(n) = 4n^4 - 41n^3 + 33n^2 + 141n + 79 < 0$ *if* $n \in [4, 8]$.

vii $P_7(n) = 25n^7 - 141n^6 + 96n^5 + 406n^4 + 261n^3 - 981n^2 - 1854n - 884 > 0$ *if* $n > 8$.

Proof. We use the following two known bounds for the positive roots of polynomials (see Ştefănesu et al., 2010). Let

$$P(n) = a_0 n^d - b_1 n^{d-m_1} - \cdots - b_k n^{d-m_k} + \sum_{j \neq m_1, \ldots, m_k} a_j n^{d-j},$$

with $a_0 > 0, b_1, \ldots, b_k > 0$ and $a_j \geq 0$ for all $j \notin \{b_1, \ldots, b_k\}$. The numbers

$$B_1(P) = \max\left\{(kb_1/a_0)^{1/m_1}, \ldots, (kb_k/a_0)^{1/m_k}\right\}$$

and

$$B_2(P) = 2 \times \max\left\{(b_1/a_0)^{1/m_1}, \ldots, (b_k/a_0)^{1/m_k}\right\}$$

are the upper bounds for the positive roots.

i Descartes' rule of signs states that P_1 has only one positive root, which is between 8 and 9, as $P_1(8) = 900$ and $P_1(9) = -23724$. Therefore, $P_1 < 0$ if $n > 8$, and positive otherwise.

ii $B_2(P_2) = 10.20$. Since the leading coefficient of P_2 is positive, $P_2 > 0$ if $n \geq 11$. As $P_2(10) = 139011 > P_2(9) = 84816 > P_2(8) = 49113 >$

$P_2(7) = 26688 > P_2(6) = 13407 > P_2(5) = 6096 > P_2(4) = 2421 > P_2(3) = 768 > 0$, then $P_2 > 0$ if $n \in [3, 10]$.

iii $B_1(P_3) = 4$ and $P_3(4) = 40 > P_3(3) = 18 > 0$. Since the leading coefficient is positive, $P_3 > 0$ if $n \geq 3$.

iv $P_4(n) = n^5 + 17n^4 - 51n^3 + 61n^2 + 10n - 6 > (17n^4 - 51n^3) + (10n - 6) \equiv P_4'(n)$. Both $17n^4 - 51n^3 > 0$ and $10n - 6 > 0$ if $n > 3$ and $P_4'(3) = 24$. Therefore, $P_4 > P_4' > 0$ if $n \geq 3$.

v $B_1(P_5) = 5.20$ and $P_5(5) = 2768 > P_5(4) = 872 > P_5(3) = 240 > 0$. Since the leading coefficient is positive, $P_5 > 0$ if $n \geq 3$.

vi Descartes' rule of signs states that P_6 has two positive roots, which are between 3 and 4, as $P_6(3) = 16$ and $P_6(4) = -429$, and 8 and 9, as $P_6(8) = -1289$ and $P_6(9) = 376$. Therefore, $P_6 < 0$ if $n \in [4, 8]$, and positive otherwise.

vii $B_2(P_7) = 11.28$ and $P_7(11) = 259001728 > P_7(10) = 122803476 > P_7(9) = 53066752 > 0$. Since the leading coefficient is positive, $P_7 > 0$ if $n > 8$.

■

Proof of Proposition 1. The marginal effect of the spillover rate on firm i's effective R&D is given by

$$\frac{\partial X_i}{\partial \beta} = 2(a-c)\Bigg(\frac{n-(n-1)\beta}{C^2}\left(\frac{\partial A}{\partial \beta}C - A\frac{\partial C}{\partial \beta}\right) - (n-1)\frac{A}{C}$$

$$+\frac{\beta(n-1)(n-\beta-(n-2)\bar{\beta})}{C^2}\left(\frac{\partial D}{\partial \beta}C - D\frac{\partial C}{\partial \beta}\right)$$

$$+(n-1)(n-2\beta-(n-2)\bar{\beta})\frac{D}{C}\Bigg),$$

with

$$\frac{\partial A}{\partial \beta} = (2-2\bar{\beta})n^2 + (8\bar{\beta}-4\beta-2)n + 4\beta - 8\bar{\beta} + 2,$$

$$\frac{\partial C}{\partial \beta} = (4\bar{\beta}^2 + (8\beta - 12\beta^2 - 2\gamma - 4)\bar{\beta} + 12\beta^2 - (4\gamma+8)\beta + 6\gamma)n^3$$

$$+((48\beta^2 - 16\beta + 8\gamma + 12)\bar{\beta} - 20\bar{\beta}^2 - 16\beta^3 - 12\beta^2 + 8\beta - 4)n^2$$

$$+(32\bar{\beta}^2 - (60\beta^2 + 8\beta + 2\gamma + 12)\bar{\beta} + 32\beta^3 + (4\gamma+8)\beta - 2\gamma)n$$

$$-16\bar{\beta}^2 + (24\beta^2 + 16\beta - 12\gamma + 8)\bar{\beta} - 16\beta^3 - 8\beta + 4\gamma,$$

and

$$\frac{\partial D}{\partial \beta} = (4-4n)\beta + 4n - 2.$$

When agglomerated with the other firms, $d_{ij} = 0$ and $\beta = \bar{\beta}$, this marginal effect is non-negative if

$$\left.\frac{\partial X_i}{\partial \beta}\right|_{\beta=\bar{\beta}} = \frac{2\gamma(a-c)(n^2-1)f(n,\gamma,\bar{\beta})}{((n+1)\gamma - (2n-2)\bar{\beta}^2 + (4n-2)\bar{\beta} - 2n)E} \geq 0 \;, \qquad (7)$$

where

$$f(n,\gamma,\bar{\beta}) = (1-\bar{\beta})(2\bar{\beta}^2 - 2\bar{\beta} + \gamma)n^3 + (1-2\bar{\beta})(16\bar{\beta} - 10\bar{\beta}^2 - 6 + \gamma)n^2$$

$$+(56\bar{\beta}^2 - 34\bar{\beta}^3 - (\gamma + 30)\bar{\beta} + 6 - \gamma)n + 16\bar{\beta}^3 - 18\bar{\beta}^2 + 6\bar{\beta} - \gamma$$

and

$$E = ((2\bar{\beta}^2 - 2\bar{\beta} + \gamma)n^2 + (4\bar{\beta} - 4\bar{\beta}^2 + 2\gamma - 2)n + 2\bar{\beta}^2 - 2\bar{\beta} + \gamma)^2.$$

Clearly, both $E > 0$ and $2\gamma(a - c)(n^2 - 1) > 0$. Given Assumption 1, $(n + 1)\gamma - (2n - 2)\bar{\beta}^2 + (4n - 2)\bar{\beta} - 2n$ is also always positive. Hence, the sign of equation (7) depends on the sign of $f(n, \gamma, \bar{\beta})$.

Since $f(n, \gamma, 0) = (n-1)(n^2\gamma + 2n\gamma - 6n + \gamma) > 0$, given Assumption 1, and because $f(n, \gamma, (n-1)/n) = -4(2n-1)(n-1)(n-2)^2/n^3 < 0$, there is at least one $\hat{\beta}$ such that $f(n, \gamma, \hat{\beta}) = 0$. Furthermore, $\partial f/\partial\gamma = (n+1)^2(n-1-\bar{\beta}n) \geq 0$ iff $\bar{\beta} \leq (n - 1)/n$. As $\gamma > (n + 1)/2$,

$$\frac{\partial f}{\partial \bar{\beta}} = -6(n - 8)(n - 1)^2\bar{\beta}^2 + 4(n - 1)(2n - 1)(n - 9)\bar{\beta} - n(n + 1)^2\gamma$$

$$-2n^3 + 28n^2 - 30n + 6 > -6(n-8)(n-1)^2\bar{\beta}^2 + 4(n-1)(2n-1)(n-9)\bar{\beta} - \frac{n}{2}(n+1)^3$$

$$-2n^3 + 28n^2 - 30n + 6$$

and the discriminant of this quadratic function,

$$\Delta_1 = -3n^5 + 19n^4 + 23n^3 + 133n^2 + 132n + 36,$$

is negative for $n > 8$, as shown in Lemma 1. Then, because the leading coefficient is also negative, this implies that $\partial f/\partial\bar{\beta} < 0$ if $n > 8$. Furthermore, for $n = 8$, $\partial f/\partial\bar{\beta} < -2383 - 420\bar{\beta} < 0$. For $n \in [3, 7]$, the leading coefficient is positive and the roots of the quadratic function are given by

$$\bar{\beta} = \frac{1}{6}\frac{4n^2 - 38n \pm \sqrt{\Delta_1} + 18}{(n-8)(n-1)}.$$

The larger root is greater than 1 when $n \in [3, 7]$ and the smaller root is less than 0 when $n \in [4, 7]$. Therefore, $\partial f/\partial\bar{\beta} < 0$ if $n > 3$. For $n = 3$, $\partial f/\partial\bar{\beta} < 0$ if $\bar{\beta} \geq 0.07805$. This implies that for $n = 3$, $\hat{\beta}$ is bounded above at 2/3. By the implicit function theorem,

$$\frac{\partial\bar{\beta}}{\partial\gamma} = -\frac{\partial f/\partial\gamma}{\partial f/\partial\bar{\beta}} = \frac{8}{3}\frac{3\bar{\beta} - 2}{20\bar{\beta}^2 - 40\bar{\beta} - 8\gamma + 19} \geq 0$$

if $\bar{\beta} \in [0.07805, 2/3]$. Hence, we conclude that there is always exactly one $\hat{\beta}$, where its higher bound is $(n-1)/n$ and the lower bound is given by

$$f\left(n, \frac{n+1}{2}, \hat{\beta}\right) = -2(n-8)(n-1)^2\hat{\beta}^3 + 2(n-1)(2n-1)(n-9)\hat{\beta}^2$$

$$-\frac{1}{2}(n^4 + 7n^3 - 53n^2 + 61n - 12)\hat{\beta} + \frac{1}{2}(n-1)(n^3 + 3n^2 - 9n + 1) = 0.$$

Since

$$f\left(n, \frac{n+1}{2}, \frac{n-1}{n+1}\right) = \frac{(n-1)(n^5 + 5n^4 - 26n^3 + 178n^2 - 287n + 81)}{2(n+1)^3} > 0,$$

with $n^5 + 5n^4 - 26n^3 + 178n^2 - 287n + 81 > 0$, as shown in Lemma 1, the lower bound is greater than $(n-1)/(n+1)$. ■

Proof of Proposition 2. The first order condition of (6) with respect to β is

$$\frac{\partial \pi_i}{\partial \beta} = \frac{2(a - c + (n-(n-1)\beta)x_i^* + (n-1)(n\beta - (n-2)\bar{\beta} - 1)x_{-i}^*)^2}{(n+1)^2}$$

$$\times\Bigg((n-(n-1)\frac{\partial x_i^*}{\partial \beta} - (n-1)x_i^* + (n-1)(n\beta-(n-2)\bar{\beta}-1)\frac{\partial x_{-i}^*}{\partial \beta}$$

$$+n(n-1)x_{-i}^*\Bigg) - \gamma x_i^* \frac{\partial x_i^*}{\partial \beta}$$

with

$$\frac{\partial x_i^*}{\partial \beta} = 2(a-c)\left(\frac{n-(n-1)\beta}{C^2}\left(\frac{\partial A}{\partial \beta}C - A\frac{\partial C}{\partial \beta}\right) - (n-1)\frac{A}{C}\right)$$

and

$$\frac{\partial x_{-i}^*}{\partial \beta} = 2(a-c)\left(\frac{n-\beta-(n-2)\bar{\beta})}{C^2}\left(\frac{\partial D}{\partial \beta}C - D\frac{\partial C}{\partial \beta}\right) + \frac{D}{C}\right).$$

Agglomeration, where $d_{ij} = 0$ and $\beta = \bar{\beta}$, is an equilibrium only if the marginal profit of agglomeration spillovers is non-negative, i.e.,

$$\left.\frac{\partial \pi_i}{\partial \beta}\right|_{\beta=\bar{\beta}} = \frac{4(n-1)\gamma(a-c)^2 h(n,\gamma,\bar{\beta})}{((n+1)\gamma + (2-2n)\bar{\beta}^2 + (4n-2)\bar{\beta} - 2n)G} \geq 0 \qquad (8)$$

where

$$G = ((2\bar{\beta}^2 - 2\bar{\beta} + \gamma)n^2 - (4\bar{\beta}^2 - 4\bar{\beta} - 2\gamma - 2)n + 2\bar{\beta}^2 - 2\bar{\beta} + \gamma)^3$$

and

$$h(n, \gamma, \bar{\beta}) = (1 - \bar{\beta})(2\bar{\beta}^2 - 2\bar{\beta} + \gamma)(4\bar{\beta} - 2\bar{\beta}^2 + \gamma - 2)n^5$$

$$-(28\bar{\beta}^5 - 100\bar{\beta}^4 + 136\bar{\beta}^3 - (2\gamma + 88)\bar{\beta}^2 + (\gamma^2 + 28)\bar{\beta} - 2\gamma^2 + 2\gamma - 4)n^4$$

$$+(72\bar{\beta}^5 - 208\bar{\beta}^4 + (20\gamma + 216)\bar{\beta}^3 - (42\gamma + 96)\bar{\beta}^2 + (26\gamma + 16)\bar{\beta} + \gamma^2 - 6\gamma)n^3$$

$$-(88\bar{\beta}^5 - 184\bar{\beta}^4 + (20\gamma + 120)\bar{\beta}^3 - (10\gamma + 24)\bar{\beta}^2 + (4\gamma^2 - 10\gamma)\bar{\beta} - \gamma^2 + 4\gamma)n^2$$

$$+(52\bar{\beta}^5 - 64\bar{\beta}^4 - (20\gamma - 16)\bar{\beta}^3 + 44\bar{\beta}^2\gamma - (7\gamma^2 + 18\gamma)\bar{\beta} + 2\gamma^2 + 2\gamma)n$$

$$(2\bar{\beta}^3 - 3\bar{\beta}\gamma + 1\gamma)(2\bar{\beta} - 6\bar{\beta}^2 + \gamma).$$

Clearly, $4(n-1)\gamma(a-c)^2 > 0$. Given Assumption 1, $(n+1)\gamma + (2-2n)\bar{\beta}^2 + (4n-2)\bar{\beta} - 2n$ and G are always positive. Hence, the sign of equation (8) depends on the sign of $h(n, \gamma, \bar{\beta})$.

At the end points,

$$h(n, \gamma, 0) = (\gamma^2 - 2\gamma)n^5 + (2\gamma^2 - 2\gamma + 4)n^4 + (\gamma^2 - 6\gamma)n^3 + (\gamma^2 - 4\gamma)n^2 + (2\gamma^2 + 2\gamma)n + \gamma^2$$

$$> h(n, \gamma, 1) = \gamma^2 n^4 + (\gamma^2 - 2\gamma)n^3 - (3\gamma^2 + 4\gamma)n^2 - (5\gamma^2 - 8\gamma - 4)n - 2\gamma^2 + 10\gamma - 8 > 0,$$

given Assumption 1. $\bar{\beta} = 1$ is not the argument that minimises $h(n, \gamma, \bar{\beta})$ only if $h(n, \gamma, \bar{\beta})$ is convex downward.

The second derivative of $h(n, \gamma, \bar{\beta})$ is:

$$\frac{\partial^2 h}{\partial^2 \bar{\beta}} = 4(n-1)((20n^4 - 120n^3 + 240n^2 - 200n + 60)\bar{\beta}^3$$

$$-(48n^4 - 252n^3 + 372n^2 + 180n - 12)\bar{\beta}^2 + (36n^4 - 168n^3 + 30n^2\gamma + 156n^2 - 24n - 30\gamma)\bar{\beta}$$

$$-n^4\gamma - 8n^4 + 36n^3 - 21n^2\gamma - 12n^2 - 16n\gamma + 6\gamma) \equiv h''.$$

Note that

$$\frac{\partial h''}{\partial \gamma} = -4(n^2 - 1)(n^3 - n^2 - 30n\bar{\beta} + 22n + 30\bar{\beta} - 6) < 0,$$

since $n^3 - n^2 - 30n\bar{\beta} + 22n + 30\bar{\beta} - 6 \geq n^3 - n^2 - 8n + 24 > 0$, as established in Lemma 1. Therefore,

$$h'' < 2(n-1)(40(n-3)(n-1)^3)\bar{\beta}^3 - 24(4n^2 - 13n + 1)(n-1)^2\bar{\beta}^2$$

$$+6(n-1)(12n^3 - 39n^2 + 18n + 5)\bar{\beta} - n^5 - 17n^4 + 51n^3 - 61n^2 - 10n + 6 \equiv \bar{h}'',$$

where $\gamma = (n+1)/2$. When $n = 3$, $\bar{h}''$ becomes a quadratic function that is always negative:

$$\bar{h}'' = 768\bar{\beta}^2 + 1536\bar{\beta} - 3264 < 0.$$

$\bar{h}''$ is also negative at both end points:

$$\bar{\beta} = 0 \to \bar{h}'' = -2(n-1)(n^5 + 17n^4 - 51n^3 + 61n^2 + 10n - 6) < 0,$$

and

$$\bar{\beta} = 1 \to \bar{h}'' = -2(n-1)(n^5 + n^4 - 9n^3 - 17n^2 + 128n - 72) < 0,$$

as shown in Lemma 1. Differentiating $\bar{h}''$ with respect to $\bar{\beta}$ gives a quadratic equation,

$$\frac{\partial \bar{h}''}{\partial \bar{\beta}} = 12(n-1)^2(20(n-3)(n-1)^2\bar{\beta}^2 - 8(4n^2-13n+1)(n-1)\bar{\beta} + 12n^3 - 39n^2 + 18n + 5, \tag{9}$$

with solutions

$$\bar{\beta} = \frac{8\,n^2 - 26\,n \pm \sqrt{\Delta_2} + 2}{10\,(n-3)\,(n-1)}. \tag{10}$$

For $n \in [4, 8]$, the discriminant,

$$\Delta_2 = 4n^4 - 41n^3 + 33n^2 + 141n + 79,$$

is negative, as shown in Lemma 1, and $\bar{h}''$ has no local maximum. Since the leading coefficient of equation (9) is positive for $n > 3$, the local maximum for $n > 8$ is given by the smaller value of equation (10). At this point, $\bar{h}''$ is

$$-\frac{2(n-1)}{25(n-3)^2}(25n^7 - 141n^6 + 96n^5 - (12\sqrt{\Delta_2} - 406)n^4 + (123\sqrt{\Delta_2} + 261)n^3$$

$$-(99\sqrt{\Delta_2}+981)n^2-(423\sqrt{\Delta_2}+1854)n+\Delta_2^{\left(\frac{3}{2}\right)}-237\sqrt{\Delta_2}-884).$$

and decreasing in Δ_2, and thus it is less than

$$-\frac{2(n-1)}{25(n-3)^2}(25n^7-141n^6+96n^5+406n^4+261n^3-981n^2-1854n-884)<0,$$

given the sign of $25n^7-141n^6+96n^5+406n^4+261n^3-981n^2-1854n-884$, as shown in Lemma 1. Since $h''<\bar{h}''<0$, $h(\gamma,\bar{\beta})$ is concave and always positive, as is then the marginal profit of agglomeration spillovers. ■